ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
XIX, 2 (1987), pp. 257-264

H. TRUSZCZYŃSKA (Warszawa)

FIXED PRECISION ESTIMATION OF THE PARAMETERS OF A LINEAR REGRESSION MODEL WITH UNKNOWN COVARIANCE STRUCTURE

1. Introduction. Let y be a sequence of random variables of the form

(1) $$y_n = x_n^T \beta + \varepsilon_n,$$

where β is an unknown p-dimensional vector, x_n^T is the transpose of a known vector x_n, and ε_n is a normally distributed random variable with zero mean. The covariance structure of (ε_n) is described by an unknown sequence $\Sigma = (\Sigma_n)$, where Σ_n is the covariance matrix of the n-dimensional random variable $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)^T$. It is of interest to estimate with fixed precision the parameter $H\beta$, where H is an $(m \times p)$-matrix $(m \leqslant p)$ of rank m. This problem has been considered in the literature but usually under some additional assumptions. Albert [1] investigated the case where $\Sigma_n = \sigma^2 I_n$ (I_n denotes the $n \times n$ identity matrix), Gleser [2], Srivastava [6], [7] and Truszczyńska [8] studied the case where $\Sigma_n = \sigma^2 I_n$ and $H = I$. The case where observations may have different variances and be correlated was considered by Gołdys [3] and Zieliński [9], but only, in particular, for $y_n = m + \varepsilon_n$. The goal of this paper is to extend the results of Zieliński [9] to the general case of (y_n) given by (1).

2. A sequential procedure. Following Zieliński [9] let us suppose that for every n $(n = 1, 2, \ldots)$ we can observe simultaneously k $(k > 1)$ independent replications $y_n^{(i)}$ of y_n given by (1). Let X_n be the $(n \times p)$-matrix whose row vectors are x_j^T $(j = 1, 2, \ldots, n)$. All results of this section are proved under the assumption that the rank of X_n is equal to p. In the last section we show that this assumption can be dropped. Let us define $Y_n^{(i)}$ to be the n-dimensional column vector whose components are $y_j^{(i)}$, $j = 1, 2, \ldots, n$, and let $\hat{\beta}_n^{(i)}$ be the least square estimator of β upon the observation vector $Y_n^{(i)}$, i.e.

$$\hat{\beta}_n^{(i)} = (X_n^T X_n)^{-1} X_n Y_n^{(i)},$$

and denote the covariance matrix of $\hat{\beta}_n^{(i)}$, which is equal to

$$(X_n^T X_n)^{-1} X_n \Sigma_n X_n (X_n^T X_n)^{-1},$$

by $\Gamma_n = (\gamma_{ij}^n)$. Finally, define an estimator $\hat{\beta}_n$ of the parameter β by

$$\hat{\beta}_n = \frac{1}{k} \sum_{i=1}^{k} \hat{\beta}_n^{(i)},$$

and an estimator $\hat{\Gamma}_n$ of the unknown covariance matrix of $\hat{\beta}_n$ by

$$\hat{\Gamma}_n = \frac{1}{k-1} \sum_{i=1}^{k} (\hat{\beta}_n^{(i)} - \hat{\beta}_n)(\hat{\beta}_n^{(i)} - \hat{\beta}_n)^T.$$

Now, we are ready to construct a sequential procedure for estimating the parameter $H\beta$ with fixed precision. Let N be a random variable defined by

(2) $$N = \inf\{n \geqslant n_0 \geqslant p+1 \colon n^a \max_{1 \leqslant i \leqslant m} \sqrt{\hat{s}_{ii}^n} \leqslant d\},$$

where a is a fixed real number and $\hat{s}_{ii}^n$ is the i-th diagonal element of the matrix $H\hat{\Gamma}_n H^T$, and let

$$R_n(d) = \{\xi \colon (\xi - H\hat{\beta}_n)^T (\xi - H\hat{\beta}_n) \leqslant d^2\}.$$

The procedure looks as follows: Take N observations and use $R_N(d)$ as the confidence region.

The properties of this procedure are summarized in the following theorem:

THEOREM. *Let (y_n) be a sequence of random variables defined by* (1) *and suppose that for the sequence (X_n), of design matrices, the sequence (Σ_n) of covariance matrices of (ε_n), and for some real number b the following condition holds:*

(3) $$\lim_{n \to \infty} n^b \max_{1 \leqslant i \leqslant p} \gamma_{ii}^n = 0.$$

Furthermore, suppose that one can observe simultaneously k independent replications of y_n, where k satisfies the inequality $1/(k-1) < b/2$. Then, for every a such that $1/(k-1) < a < b/2$, the stopping rule N defined by (2) *has the following properties:*

(i) $P\{N < \infty\} = 1$,

(ii) $\mathrm{E}(N^j) < \infty$ *for every* $j = 1, 2, \ldots$,

(iii) *for every $d > 0$ and $\alpha \in (0, 1)$, there exists $n_0(\alpha) \geqslant p+1$ such that for N defined by* (2) *with $n_0 = n_0(\alpha)$ the inequality $P\{H\beta \in R_N(d)\} \geqslant 1-\alpha$ holds.*

Proof. (i) It suffices to show that $\lim_{n\to\infty} P\{N>n\}=0$. We have

$$P\{N>n\} \leqslant P\{n^a \max_{1\leqslant i\leqslant m} \sqrt{\hat{s}_{ii}^n} > d\} = P\left\{\hat{s}_{11}^n > \frac{d^2}{n^{2a}} \text{ or } \ldots \text{ or } \hat{s}_{mm}^n > \frac{d^2}{n^{2a}}\right\}$$

$$\leqslant \sum_{i=1}^{m} P\left\{\hat{s}_{ii}^n > \frac{d^2}{n^{2a}}\right\} \leqslant \sum_{i=1}^{m} \frac{n^{2a}\,\mathrm{E}(\hat{s}_{ii}^n)}{d^2}.$$

The last inequality follows from the Chebyshev inequality. Since $2a<b$, we have

$$\sum_{i=1}^{m} \frac{n^{2a}\,\mathrm{E}(\hat{s}_{ii}^n)}{d^2} = \frac{n^{2a}}{d^2}\operatorname{trace}(H\Gamma_n H^T) \leqslant \frac{mn^{2a}}{d^2}\|H\|^2\|\Gamma_n\| \leqslant \frac{mn^b}{d^2}\|H\|^2 p^2 \max_{1\leqslant i\leqslant p} \gamma_{ii}^n$$

($\|\cdot\|$ denotes the Euclidian norm). Condition (3) and the obtained inequalities imply that $\lim_{n\to\infty} P\{N>n\}=0$, as required.

(ii) We have

$$\mathrm{E}(N^j) = \sum_{n=n_0}^{\infty} n^j P\{N=n\} = \sum_{n=n_0}^{\infty} n^j P\left\{\max_{1\leqslant i\leqslant m} \sqrt{\hat{s}_{ii}^{n-1}} > \frac{d^2}{(n-1)^{2a}}\right\}$$

$$\leqslant \sum_{n=n_0-1}^{\infty} (n+1)^j P\left\{\max_{1\leqslant i\leqslant m} \hat{s}_{ii}^n > \frac{d^2}{n^{2a}}\right\} \leqslant \sum_{n=n_0-1}^{\infty} (n+1)^j \sum_{i=1}^{m} P\left\{\hat{s}_{ii}^n > \frac{d^2}{n^{2a}}\right\}.$$

Since $\hat{s}_{ii}^n/s_{ii}^n$, where s_{ii}^n denotes the i-th diagonal element of $H\Gamma_n H^T$, has the χ_{k-1}^2-distribution, we obtain

$$\sum_{i=1}^{m} P\left\{\hat{s}_{ii}^n > \frac{d^2}{n^{2a}}\right\} \leqslant \sum_{i=1}^{m} P\left\{\chi_{k-1}^2 > \frac{d^2}{n^{2a}\max_{1\leqslant i\leqslant m} s_{ii}^n}\right\} \leqslant mc_k \exp\left\{-\frac{d^2}{n^{2a}\max_{1\leqslant i\leqslant m} s_{ii}^n}\right\}.$$

The last inequality follows from the inequality $P\{\chi_k^2 > B\} \leqslant c_k e^{-B/4}$, which is easy to prove (here c_k is a constant depending on k only). Thus we have

$$\mathrm{E}(N^j) \leqslant mc_k \sum_{n=1}^{\infty} n^j \exp\left\{-\frac{d^2}{n^{2a}\max_{1\leqslant i\leqslant m} s_{ii}^n}\right\}. \tag{4}$$

The condition (3) implies that there exists a constant M such that

$$\max_{1\leqslant i\leqslant m} s_{ii}^n < \frac{M}{n^b}. \tag{5}$$

From (4) and (5) we get

$$\mathrm{E}(N^j) \leqslant mc_k \sum_{n=1}^{\infty} n^j \exp\left\{-\frac{d^2 n^{b-2a}}{M}\right\}.$$

Since $2a<b$, the above series is convergent for every $j=1,2,\ldots$

(iii) Let us denote the parametric function $H\beta$ by u, the least square estimator of $H\beta$ based on n observations by $\hat{u}_n$, and their components by u_i and $\hat{u}_{in}$, respectively. In other words,

$$H\beta = u = (u_1, \ldots, u_m)^T, \quad H\hat{\beta} = \hat{u}_n = (\hat{u}_{1n}, \ldots, \hat{u}_{mn})^T.$$

Let C_n be an m-dimensional cube defined in the following way:

$$C_n = \{z \in R^m: |\hat{u}_{in} - z_i| < \sqrt{d^2/m} \text{ for every } i = 1, \ldots, m\}.$$

Since $C_N \subset R_N$, we have

$$\begin{aligned}
P\{H\beta \notin R_N\} &< P\{H\beta \notin C_N\} \\
&= P\{|\hat{u}_{1N} - u_1| > \sqrt{d^2/m} \text{ or } \ldots \text{ or } |\hat{u}_{mN} - u_m| > \sqrt{d^2/m}\} \\
&\leqslant \sum_{n=n_0}^{\infty} P\{|\hat{u}_{1n} - u_1| > \sqrt{d^2/m} \text{ or } \ldots \text{ or } |\hat{u}_{mn} - u_m| > \sqrt{d^2/m} \\
&\qquad \text{and } n^a \sqrt{\hat{s}_{11}^n} < d \text{ and } \ldots \text{ and } n^a \sqrt{\hat{s}_{mm}^n} < d\} \\
&\leqslant \sum_{n=n_0}^{\infty} \sum_{i=1}^{m} P\{|\hat{u}_{in} - u_i| > \sqrt{d^2/m} \text{ and } n^a \sqrt{\hat{s}_{ii}^n} < d\} \\
&\leqslant \sum_{n=n_0}^{\infty} \sum_{i=1}^{m} P\left\{\frac{\sqrt{k}\,|\hat{u}_{in} - u_i|}{\sqrt{\hat{s}_{ii}^n}} > n^a \sqrt{k/m}\right\} \\
&= m \sum_{n=n_0}^{\infty} P\{|t_{k-1}| > n^a \sqrt{k/m}\}.
\end{aligned}$$

Using the approximation for tail areas given by Grass and Hosmer in [4] we get

$$P\{|t_{k-1}| > n^a \sqrt{k/m}\} \leqslant \frac{2\Gamma(k/2)(k-1)^2 m^{(k-1)/2}}{\sqrt{\pi}\,\Gamma((k-1)/2)(5(k-1)+(k-1)^2+2)} n^{-a(k-1)}.$$

Since $1/(k-1) < a$, the series

$$\sum_{n=n_0}^{\infty} P\{|t_{k-1}| > n^a \sqrt{k/m}\}$$

is convergent. Thus there exists a constant $n_0(\alpha)$ such that

$$m \sum_{n=n_0(\alpha)}^{\infty} P\{|t_{k-1}| > n^a \sqrt{k/m}\} < \alpha.$$

3. Condition (3) for some special design matrices. To apply the procedure described in Section 2, design matrices (X_n) and covariance matrices (Σ_n) have to satisfy (3). We shall show that if the sequence of design matrices (X_n) fulfils certain conditions, then the class of covariance matrices (Σ_n) which, together with (X_n), satisfies condition (3) is quite large.

3.1. Let $K = \{\Sigma = (\sigma_{ij})_{i,j=1,2,\ldots}$: there exists M such that $\sigma_{ii} < M$ for every $i = 1, 2, \ldots\}$.

PROPOSITION 1. *Assume that a sequence of design matrices* (X_n) *satisfies the following conditions*:

(a) *there exists* $\alpha > 1$ *such that*

$$\limsup_{n\to\infty} \frac{n^\alpha}{\lambda_{\min}(X_n^T X_n)} < \infty,$$

(b) $\limsup_{n\to\infty} \dfrac{\lambda_{\max}(X_n^T X_n)}{\lambda_{\min}(X_n^T X_n)} < \infty.$

Then for (X_n) *and for every* $\Sigma \in K$ *condition* (3) *is fulfilled whenever* $b \in (0, \alpha - 1)$.

Proof. Let

$$M_1 = \limsup_{n\to\infty} \frac{n^\alpha}{\lambda_{\min}(X_n^T X_n)} \quad \text{and} \quad M_2 = \limsup_{n\to\infty} \frac{\lambda_{\max}(X_n^T X_n)}{\lambda_{\min}(X_n^T X_n)}.$$

For $\Sigma \in K$ we have

$$\gamma_{ii}^n \leqslant \|(X_n^T X_n)^{-1}\|^2 \|X_n\|^2 \|\Sigma_n\| \leqslant p\lambda_{\max}^2((X_n^T X_n)^{-1}) \operatorname{trace}(X_n^T X_n) \|\Sigma_n\|$$
$$\leqslant p^2 Mn \frac{\lambda_{\max}(X_n^T X_n)}{\lambda_{\min}^2(X_n^T X_n)} \leqslant p^2 MM_2 \frac{n}{\lambda_{\min}(X_n^T X_n)} \leqslant p^2 MM_1 M_2 \frac{1}{n^{\alpha-1}}.$$

Hence

$$\lim_{n\to\infty} n^b \max_{1\leqslant i\leqslant p} \gamma_{ii}^n \leqslant \lim_{n\to\infty} \left(p^2 MM_1 M_2 \frac{n^b}{n^{\alpha-1}}\right) = 0,$$

since $b < \alpha - 1$.

3.2. Let γ be a positive constant and let

$$K = \{\Sigma = (\sigma_{ij})_{i,j=1,2,\ldots}: \limsup_{n\to\infty} \Big(\sum_{i,j=1}^{m} |\sigma_{ij}|/n^{2-\gamma}\Big) < \infty\}.$$

PROPOSITION 2. *Assume that a sequence of design matrices satisfies the following conditions*:

(a) $\lim_{n\to\infty} \dfrac{n}{\lambda_{\min}(X_n^T X_n)} < \infty,$

(b) $\limsup_{n\to\infty} \dfrac{\lambda_{\max}(X_n^T X_n)}{\lambda_{\min}(X_n^T X_n)} < \infty.$

Then for (X_n) *and for every* $\Sigma \in K$ *condition* (3) *is fulfilled whenever* $b \in (0, \gamma)$.

Proof. Let J_p denote the matrix of ones of order p and let

$$M_1 = \lim_{n\to\infty} \frac{n}{\lambda_{\min}(X_n^T X_n)} \quad \text{and} \quad M_2 = \limsup_{n\to\infty} \sum_{i,j=1}^{n} |\sigma_{ij}|/n^{2-\gamma}.$$

Conditions (a) and (b) imply that there exists a constant $c > 0$ such that, for every (i, j) and for every n, $|(X_n)_{ij}| < c$. Hence

$$\gamma_{ii}^n \leqslant \frac{p}{\lambda_{\min}^2(X_n^T X_n)} \|X_n^T \Sigma_n X_n\| \leqslant \frac{pc}{\lambda_{\min}^2(X_n^T X_n)} \sum_{i,j=1}^{n} |\sigma_{ij}| \, \|J_p\|$$
$$\leqslant p^2 c M_2 \frac{n^{2-\gamma}}{\lambda_{\min}^2(X_n^T X_n)}.$$

Consequently,

$$\lim_{n\to\infty} n^b \max_{1\leqslant i\leqslant p} \gamma_{ii}^n \leqslant p^2 c M_2 M_1 \lim_{n\to\infty} n^{b-\gamma} = 0,$$

since $b < \gamma$.

3.3. Let $K = \{\Sigma = (\sigma_{ij})_{i,j=1,2,\ldots}\colon \sigma_{ij} = 0$ for $|i-j| > r$ and there exists M such that, for every i, $\sigma_{ii} < M\}$ (Σ is a covariance structure in the case where only r nearest observations are correlated).

PROPOSITION 3. *Assume that a sequence of design matrices satisfies the following conditions*:

(a) $\limsup_{n\to\infty} \dfrac{n^\alpha}{\lambda_{\min}(X_n^T X_n)} < \infty$ *for some* $\alpha > 0$,

(b) $\limsup_{n\to\infty} \dfrac{\lambda_{\max}(X_n^T X_n)}{\lambda_{\min}(X_n^T X_n)} < \infty$.

Then for (X_n) *and for every* $\Sigma \in K$ *condition* (3) *is fulfilled whenever* $b < \alpha$.

Proof. Let

$$M_1 = \limsup_{n\to\infty} \frac{n^\alpha}{\lambda_{\min}(X_n^T X_n)}, \quad M_2 = \limsup_{n\to\infty} \frac{\lambda_{\max}(X_n^T X_n)}{\lambda_{\min}(X_n^T X_n)}.$$

For $\Sigma \in K$ we have

$$\gamma_{ii}^n \leqslant \lambda_{\max}\big((X_n^T X_n)^{-1} X_n^T \Sigma_n X_n (X_n^T X_n)^{-1}\big)$$
$$\leqslant \frac{1}{\lambda_{\min}^2(X_n^T X_n)} \lambda_{\max}(X_n^T X_n) \lambda_{\max}(\Sigma_n) \leqslant M_1 M_2 \frac{\lambda_{\max}(\Sigma_n)}{n^\alpha}.$$

Using the inequality

$$\lambda_{\max}(A) \leqslant \max_i \sum_{i\neq j} |a_{ij}|$$

we obtain

$$\lim_{n\to\infty} n^b \max_{1\leqslant i\leqslant p} \gamma_{ii}^n \leqslant 2r M_1 M_2 M \lim_{n\to\infty} n^{b-\alpha} = 0,$$

since $b < \alpha$.

4. Generalization. In this section we shall show that the assumption that the rank of design matrices (X_n) is equal to the number of parameters can be dropped.

Let us suppose that there exists n_0 such that, for every $n \geqslant n_0$, $\operatorname{rank}(X_n) = j$ $(0 < j < p)$. Let $(X_n^T X_n)^-$ be a g-inverse of $(X_n^T X_n)$ and let $H\beta$ be an arrangement of m linearly independent estimable parametric functions $(m \leqslant j)$. Denote by $\mathscr{R}(X_n^T X_n)$ the vector space spanned by the columns of $X_n^T X_n$. For every $n \geqslant n_0$, $\mathscr{R}(X_n^T X_n) = \mathscr{R}(X_n^T X_n)$. For a vector $h \in \mathscr{R}(X_n^T X_n)$ define $\gamma_n(h)$ by

$$\gamma_n(h) = h^T (X_n^T X_n)^- X_n^T \Sigma_n X_n (X_n^T X_n)^- h.$$

It is well known [5] that $\gamma_n(h)$ has the same value for every g-inverse of $X_n^T X_n$. Finally, define $\hat{\beta}_n$ as

$$\hat{\beta}_n = \frac{1}{k} \sum_{i=1}^{k} (X_n^T X_n)^- X_n^T Y_n^{(i)}$$

and replace condition (3) by

(3′) $$\lim_{n\to\infty} n^b \gamma_n(h) = 0 \quad \text{for every } h \in \mathscr{R}(X_{n_0}^T X_{n_0}).$$

Subject to these changes the Theorem of Section 2 remains valid.

Remark. *Propositions 1, 2 and 3 of Section 3 remain true if in the assumptions of these propositions $\lambda_{\min}(X_n^T X_n)$, which now does not exist, is replaced by the smallest, greater than 0, eigenvalue of $X_n^T X_n$.*

We shall prove the remark only for Proposition 1. Other cases can be dealt with similarly. Since $\gamma_n(h)$ has the same value for any g-inverse of $X_n^T X_n$, it suffices to show that, for a certain sequence of g-inverses of $X_n^T X_n$,

$$\lim_{n\to\infty} n^b \gamma_n(h) = 0.$$

Let $(X_n^T X_n)^- = (X_n^T X_n + \hat{\lambda}(X_n^T X_n) B^T B)^{-1}$, where B is a $((p-j)\times p)$-matrix whose row vectors are orthogonal to the row vectors of X_{n_0}, and $\hat{\lambda}(A)$ denotes the smallest, greater than 0, eigenvalue of A. Since

$$\lambda_{\min}(X_n^T X_n + \hat{\lambda}(X_n^T X_n) B^T B)$$
$$= \min\{\hat{\lambda}(X_n^T X_n),\ \hat{\lambda}(X_n^T X_n),\ \hat{\lambda}(B^T B)\} = \min\{1,\ \hat{\lambda}(B^T B)\}\, \hat{\lambda}(X_n^T X_n)$$

and since $M_1 < \infty$ and $M_2 < \infty$ (notation as in the proof of Proposition 1), for every h and $\Sigma \in K$ we have

$$\lim_{n\to\infty} n^b \gamma_n(h) \leqslant \|h\|^2 \lim_{n\to\infty} \|(X_n^T X_n)^- X_n^T \Sigma_n X_n (X_n^T X_n)^- \|$$
$$\leqslant \|h\|^2 \lim_{n\to\infty} \left(\|(X_n^T X_n)^-\|^2 \|X_n\|^2 \|\Sigma_n\|\right)$$
$$\leqslant \|h\|^2 p^2 \lim_{n\to\infty} \left(\frac{\lambda_{\max}(X_n^T X_n)}{\lambda_{\min}(X_n^T X_n + \hat{\lambda}(X_n^T X_n) B^T B)} \|\Sigma_n\|\right) = 0$$

since $b < \alpha - 1$.

Acknowledgement. The author is grateful to Ryszard Zieliński for his help during the preparation of this paper.

References

[1] A. Albert, *Fixed size confidence ellipsoids for linear regression parameters*, Ann. Math. Statist. 37 (1966), pp. 1602–1630.

[2] L. J. Gleser, *On the asymptotic theory of fixed-size sequential confidence bound for linear regression parameters*, ibidem 36 (1965), pp. 463–467.

[3] B. Gołdys, *Fixed precision estimation of the mean of a Gaussian sequence*, Doctoral thesis, Technical University, Warszawa 1981.

[4] A. J. Grass and D. W. Hosmer, Jr., *Approximation tail areas of probability distributions*, Ann. Statist. 6 (1978), pp. 1352–1359.

[5] C. R. Rao, *Linear statistical inference and its applications*, J. Wiley, New York 1973.

[6] M. S. Srivastava, *On fixed-width confidence bound for regression parameters and mean vector*, J. Roy. Statist. Soc. B 29 (1967), pp. 132–140.

[7] — *On fixed-width confidence bound for regression parameters*, Ann. Math. Statist. 42 (1971), pp. 1403–1411.

[8] H. Truszczyńska, *Fixed precision estimation of linear regression parameters*, Matematyka Stosowana 23 (1984), pp. 65–76.

[9] R. Zieliński, *Fixed precision estimation of the mean of a Gaussian sequence with unknown covariance structure*, Proc. 6th Internat. Conference Math. Statist. and Prob. Theory, Wisła 1978.

AGRICULTURAL ACADEMY
00-950 WARSZAWA

Received on 1984.12.19;
revised version on 1986.01.23